Systematic Botany (2005), 30(3): pp. 481–493

Pattern of Variation and Systematics of *Nymphaea odorata*: II. Sequence Information from ITS and *trnL-trnF*

KRISTI WOODS,[1,4,5] KHIDIR W. HILU,[1] THOMAS BORSCH,[2] and JOHN H. WIERSEMA[3]

[1]Virginia Polytechnic Institute and State University, 2119 Derring, Blacksburg, Virginia 24061;
[2]Nees-Instiut für Biodiversität der Pflanzen, Friedrich-Wilhelms-Universität Bonn, Meckenheimer Allee 170, 53115 Bonn, Germany
[3]United States Department of Agriculture, Agricultural Research Service, Building 011A, BARC-West, Beltsville, Maryland 20705-2350;
[4]Present address: Novozymes Biologicals, Inc., 5400 Corporate Circle, Salem, Virginia 24153;
[5]Author for Correspondence (kyw@novozymes.com)

Communicating Editor: Gregory M. Plunkett

ABSTRACT. Sequence data from the nuclear internal transcribed spacer (ITS) and the plastid *trnL-trnF* regions were used to assess relationships among populations of *N. odorata* across its North American range, and to evaluate whether subsp. *odorata* and subsp. *tuberosa* form distinct taxonomic units. *Nymphaea mexicana* was included because of suspected hybridization with *N. odorata*. The *trnL-trnF* region provided a single informative site in *N. odorata*. In contrast, the ITS region was more variable. Phylogenetic analysis of ITS data supports the monophyly of the two species. Within *N. odorata*, two clades were resolved largely representing subsp. *odorata* and subsp. *tuberosa*, although a few individuals appeared outside the respective clades. Polymorphic sites were detected in ITS, indicating possible hybridization between the subspecies. The geographic location of these hybrids suggests a possible hybrid zone. Overall, molecular evidence supports the segregation of subsp. *odorata* and subsp. *tuberosa*, with limited gene flow between them.

Nymphaea L. is the most diverse genus in the waterlily family Nymphaeaceae. Borsch et al. (1998) identified five major clades within *Nymphaea* based on sequence variation of plastid *trnT-trnF* and *matK,* and nuclear internal transcribed spacer (ITS) regions. Among these is one highly supported clade consisting of the temperate subgenus *Nymphaea*. Within this temperate clade, either *N. odorata* Aiton and *N. mexicana* Zucc. form a clade sister to the rest, or *N. mexicana* is sister to *N. odorata* plus remaining members of subg. *Nymphaea* (Borsch 2000; Borsch et al. unpubl. data). *Nymphaea odorata* and *N. mexicana* are confined to the Americas and overlap in distribution in Florida, Georgia, Alabama, Louisiana, and Texas, where putative hybrids have been reported (Ward 1977; Wiersema and Hellquist 1997). Despite possible gene flow, the two species remain morphologically distinct. The yellow petals and prominent stolons easily separate *Nymphaea mexicana* from *N. odorata*, which has white petals and lacks stolons.

Conflicting views exist on the taxonomy of *N. odorata*. In some floristic treatments, the species has been split into two, *N. odorata* and *N. tuberosa* Paine (Correll and Correll 1975; McGregor et al. 1986). In others, these two taxa are merged under *N. odorata* (Voss 1985; Gleason and Cronquist 1991), without recognition of *N. tuberosa* at any rank. More recently, two subspecies of *N. odorata*, subsp. *odorata* and subsp. *tuberosa* (Paine) Wiersema & Hellquist have been recognized primarily on the basis of petiole stripes, color of leaf blade undersurface, and seed size (Wiersema and Hellquist 1994; Wiersema and Hellquist 1997; Crow and Hellquist 2000). In subsp. *odorata*, the petiole lacks stripes, the abaxial surface of the leaf blade is reddish-purple, and the seeds are 1.5–2.5 mm long. In contrast, subsp. *tuberosa* has a striped petiole, an abaxially green leaf blade, and 2.8–4.5 mm long seeds. Differences in geographic distribution are also evident (see Woods et al. 2005, Fig. 1). However, suspected intersubspecific hybrids can be difficult to identify to either subspecies (Wiersema and Hellquist 1997).

The goal of this study was to employ a phylogenetic approach using DNA sequence data from the nuclear ITS and chloroplast *trnL-trnF* regions to analyze the evolution and diversification of *N. odorata* and to evaluate competing systems of classification. This study complements an analysis of variation in morphological characters and inter-simple sequence repeats (ISSR) markers in *N. odorata* (Woods et al. 2005).

MATERIALS AND METHODS

Material. Forty-seven samples from across the North American ranges of *N. mexicana* and *N. odorata* and one sample each of *Nymphaea ampla* (Salisb.) DC and *N. elegans* Hook. as outgroups were used for analyses of the ITS region (Table 1). Two samples collected from Florida (KN20 and KN22) were obtained late in the season from deep water and thus lacked floral parts and rhizomes. Based on leaf morphology, these samples were tentatively classified as *N. cf. mexicana.* Since *trnL-trnF* provides only one variable site and thus is not effective in phylogenetic reconstruction within *N. odorata*, certain samples were chosen for *trnL-trnF* sequencing based on the following criteria: samples (1) considered representatives for each subspecies (defined by having typical morphological characteristics and ITS sequence of the subspecies), (2) appeared in unexpected positions in the ITS tree, or (3) identified as *N. cf. mexicana.* Identification of samples to the subspecies level follows the concept of Wiersema and Hellquist (1997). The geo-

FIG. 1. Distribution of samples collected during field trips and used for DNA analysis.

graphic distribution of the samples used in this study is mapped in Fig. 1.

DNA Extraction, Amplification, and Sequencing. DNA was extracted from silica gel-dried or frozen leaf tissue of individual plants using a modified CTAB protocol (Borsch et al. 2003) that followed the miniprep procedure of Liang and Hilu (1996). ITS and *trnL-trnF* amplification followed Borsch (2000) and Borsch et al. (2003), respectively. The ITS region (ITS1 + 5.8S + ITS2) was amplified and sequenced using primers ITS 4 and ITS 5 (White et al. 1990). Both the forward and reverse strand of the ITS were sequenced. Primers c and f (Taberlet et al. 1991) were used to amplify and sequence the *trnL-trnF* region. The amplified products were separated on a 1.5% agarose gel, excised and column-cleaned using the qiaquick gel extraction kit (QIAGEN, Inc., Valencia, California). The purified PCR products were sequenced with the amplification primers utilizing the ABI Prism™ BigDye Terminator cycle sequencing Ready reaction kit (Perkin Elmer, Norwalk, Connecticut). Samples were then resolved on an ABI 377 Automated DNA Sequencer, and resulting chromatograms were manually edited using EditView version 1.0.1 (Applied Biosystems, Foster City, California).

Sequence Alignment and Phylogenetic Analysis. *Nymphaea ampla* and *N. elegans* of subgenus *Brachyceras* were used as outgroups because they are members of a clade sister to subg. *Nymphaea* (Borsch 2000). ITS sequences were aligned manually using QuickAlign (Müller 2002), and the final alignment was deposited in TreeBase (study accession number = S1163, matrix accession number = M2000). The data were analyzed with maximum parsimony (MP) and maximum likelihood (ML) methods. For parsimony analyses, heuristic searches were employed with all characters equally weighted, their states unordered, and gaps treated as missing data. Polymorphic sites were treated using the multistate taxa/polymorphism option (Swofford 2000). The search options consisted of random sequence addition for 500 replicates, holding 500 trees, using TBR branch swapping, MULPARS on, and steepest descent off. The resulting trees were used to compute a strict consensus. Bootstrap (BS) values (Felsenstein 1985) based on 500 replicates were calculated as measures of support for individual clades, following the same search conditions as the parsimony analysis. In addition, a maximum likelihood analysis was performed assuming a general time reversible model (GTR), and a rate variation among sites following a gamma distribution (four categories represented by mean). According to the Akaike Information Criterion (AIC, Akaike 1974) GTR+G was chosen as the model that best fit the data by Modeltest v3.06 (Posada and Crandall 1998), employing the windows front-end (Patti 2002). The proposed settings by Modeltest v3.06 [Base = (0.2136 0.2456 0.2771), Nst = 6, Rmat = (0.5479 1.3200 1.0370 0.2689 2.4121), Shape = 0.3407] were executed in *win*PAUP 4.0b10. All analyses were performed in PAUP* version 4.0b10 (Swofford 2002).

Analysis of Polymorphic Sites. Polymorphic sites, where overlapping peaks appear in the pherogram, were found in the ITS region and are attributed to the presence of different ITS alleles in those individual plants. Although the pherograms had minimum background, the presence of a polymorphism was confirmed by three methods: (1) presence of the same pairs of nucleotides in the forward and reverse primer sequence, (2) re-amplification and sequencing of the sample, and 3) selectively sequencing a 1:1 ratio of PCR mixture from pairs of samples that possess different sets of those polymorphic sites. These steps were taken to ensure that polymorphisms were not due to sample contamination or sequencing error.

To elucidate potential relationships among samples possessing polymorphic sites, a separate data matrix was constructed for

TABLE 1. Samples used along with their geographic origin and sources of the material. Samples cited as *N. odorata* s.l. without a subspecies affiliation were identified as putative hybrids between the subspecies. Data are presented in the following sequence: *Nymphaea* species, Sample number, Geographic origin, Voucher information, GenBank Numbers.

N. elegans, 6N, Florida, *Borsch and Wilde 3084* (BONN), AY771811

N. ampla, 100N, Mexico, *Novelo R., A. et al. 1295* (MEXU), AY771812

N. mexicana, 69N, Florida, *Borsch & Summers 3227* (BONN), AY771813; KN8, Texas, *Woods & Borsch 0701* (VPI, BONN), AY771814; KN9, Louisiana, *Woods & Borsch 1101* (VPI, BONN), AY7718151; KN21, Mexico, *Novelo et al. 1343* (MEXU), AY771816

N. cf. mexicana, KN20, Florida, *Borsch & Summers 3213* (FR), AY771817; KN22, Florida, *Borsch & Summers 3214* (FR), AY771818

N. odorata subsp. *odorata*, 5N, Maryland, *Borsch, Hilu & Wiersema 2361* (VPI, BONN), AY771819; 11N, Florida, *Borsch & Wilde 3128* (FR), AY771820; 42N, Florida, *Borsch & Wilde 3099* (FR), AY771829; 33N, Florida, *Borsch & Wilde 3101* (FR), AY771821; 34N, Florida, *Borsch & Wilde 3125* (FR), AY771822; 35N, Florida, *Borsch & Wilde 3127* (FR), AY771823; 37N, Georgia, *Borsch & Wilde 3131* (FR), AY771824; 38N, Georgia, *Borsch & Wilde 3133* (FR), AY771825; 39N, Georgia, *Borsch & Wilde 3134* (FR), AY771826; 40N, Georgia, *Borsch & Wilde 3135* (FR), AY771827; 41N, Georgia, *Borsch & Wilde 3136* (FR), AY771828; 50N, Georgia, *Borsch & Wilde 3132* (FR), AY771830; KN7, Michigan, *Borsch, Wiersema & Hellquist 3398* (VPI, BONN), AY771831; KN10, Texas, *Woods & Borsch 0801* (VPI, BONN), AY771832; KN11, Louisiana, *Woods & Borsch 0901* (VPI, BONN), AY771833; KN12, Louisiana, *Woods & Borsch 1001* (VPI, BONN), AY771834; KN18, South Carolina, *Woods & Wiersema 0601* (VPI), AY771836; KN19, Virginia, *Woods 1201* (VPI), AY771837; KN23, Vermont, *Borsch, Wiersema & Hellquist 3331* (VPI, BONN), AY771838; KN24, North Carolina, *Woods 1401* (VPI), AY771839; KN26, Tennessee, *Woods & Neves 1701* (VPI), AY771841; KN27, Vermont, *Borsch, Wiersema & Hellquist 3322* (VPI, BONN), AY771842; KN29, Vermont, *Borsch, Wiersema & Hellquist 3330* (VPI, BONN), AY771843; KN30, Florida, *Borsch & Summers 3215* (FR), AY771844; KN32, Vermont, *Borsch, Wiersema & Hellquist 3323* (VPI, BONN), AY771845; KN33, Vermont, *Borsch, Wiersema & Hellquist 3324* (VPI, BONN), AY771846; KN16, Delaware, *Woods & Wiersema 0401* (VPI), AY771835; KN25, Tennessee, *Woods & Neves 1501* (VPI), AY771840; KN37, Virginia, *Woods 1301* (VPI), AY771855

N. odorata subsp. *tuberosa*, KN28, Vermont, *Borsch, Wiersema & Hellquist 3329* (VPI, BONN), AY771853; KN5, Wisconsin, *Borsch, Wiersema & Hellquist 3396* (VPI, BONN), AY771849; 1N, New York, *Borsch 3156* (FR), AY771847; KN6, Manitoba, *Borsch, Wiersema & Hellquist 3389* (BONN), AY771848; KN14, Michigan, *Woods & Wiersema 0201* (VPI), AY771850; KN15, Ohio, *Woods & Wiersema 0301* (VPI), AY771851; KN13, Pennsylvania, *Woods & Wiersema 0101* (VPI), AY771852; KN31, Vermont, *Borsch, Wiersema & Hellquist 3325* (VPI, BONN), AY771854; KN17, Ohio, *Wiersema 2384* (VPI), AY771859

N. odorata s. l., KN1, Michigan, *Borsch & Wiersema 3399* (VPI), AY771856; KN3, Michigan, *Borsch & Wiersema 3401* (VPI), AY771857; KN4, Michigan, *Borsch & Wiersema 3402* (VPI), AY771858

these sites. In this matrix, samples were considered operational taxonomic units and polymorphic sites as characters. Outgroups (*N. elegans* and *N. ampla*) were not included in this matrix because gene flow between them and *N. odorata* and *N. mexicana* is not possible (Borsch 2000). At each potentially polymorphic site in the alignment, both of the two nucleotides were scored as an independent bistate character, presence (1) and absence (0). Therefore, for a particular site, a sample that lacks that particular polymorphism will have a score of 1, 0 or 0, 1 depending on the nucleotide present at that site, whereas a sample that possesses that polymorphic site will be scored as 1, 1 (Appendix 1). A simple matching similarity matrix (Sokal and Michener 1958) was generated from the raw matrix and used in a phenetic analysis with the sequential, agglomerative, hierarchal, and nested (SAHN) clustering, principal coordinate analysis (PCOA), and minimum spanning tree (MST); the latter was superimposed on the PCOA. NTSYS-pc package of computer programs version 2.02k was used for these analyses (Rohlf 1998).

RESULTS

ITS Region. The alignment of the ITS region is 712 bp long and required the insertion of 35 gaps, varying in length from one to eleven bp. These gaps were found primarily between the outgroup and the ingroup taxa or between *N. odorata* and *N. mexicana*. Only a single gap was found among samples of *N. odorata*, where samples 39N and KN10 lacked a T at position 687, which was present in all other ingroup samples. The ITS region (ITS1 + 5.8S + ITS2) is 656 bp in *N. mexicana* and 646–647 bp in *N. odorata*. There are 163 (25%) variable characters, of which 133 (82%) are potentially parsimony informative. Excluding outgroups, there are 44 potentially parsimony informative characters, including 12 between subsp. *odorata* and subsp. *tuberosa*.

Parsimony analysis (Fig. 2) shows samples of *N. mexicana* (plus *N. cf. mexicana*) and those of *N. odorata* each in well-supported clades (100% and 99% BS, respectively). The consensus tree (marked in Fig. 2) resolved a weakly (65% BS) supported clade that represents samples of subsp. *odorata* except for one sample of subsp. *tuberosa*. Remaining samples appeared in basal polytomies. In the bootstrap tree (shown in Fig. 2), eight of the samples in the polytomy appear in a clade that was not resolved in the consensus tree.

The overall ML tree topology (Lscore-1966.38181; Fig. 3) was similar to that of MP in that *N. mexicana* and *N. odorata* form two distinct lineages; however, resolution was improved for the two subspecies of *N. odorata*. The polytomy at the base of the *N. odorata* clade in the MP tree (Fig. 2) was resolved by ML into a dichotomy, with one clade comprised primarily of subsp. *tuberosa* populations and another that includes subsp. *odorata* (except for KN28) in one subclade and a mixture of subsp. *odorata* and subsp. *tuberosa* populations in another. Components of the latter clade appeared unresolved in the MP.

trnL-trnF *Region.* Borsch (2000) has shown that in *Nymphaea*, position 375 of the *trnL* intron is variable (T or G), and may represent a molecular marker that discriminates between the two subspecies of *N. odorata*. This base is a T in *N. mexicana* and subsp. *odorata*, and

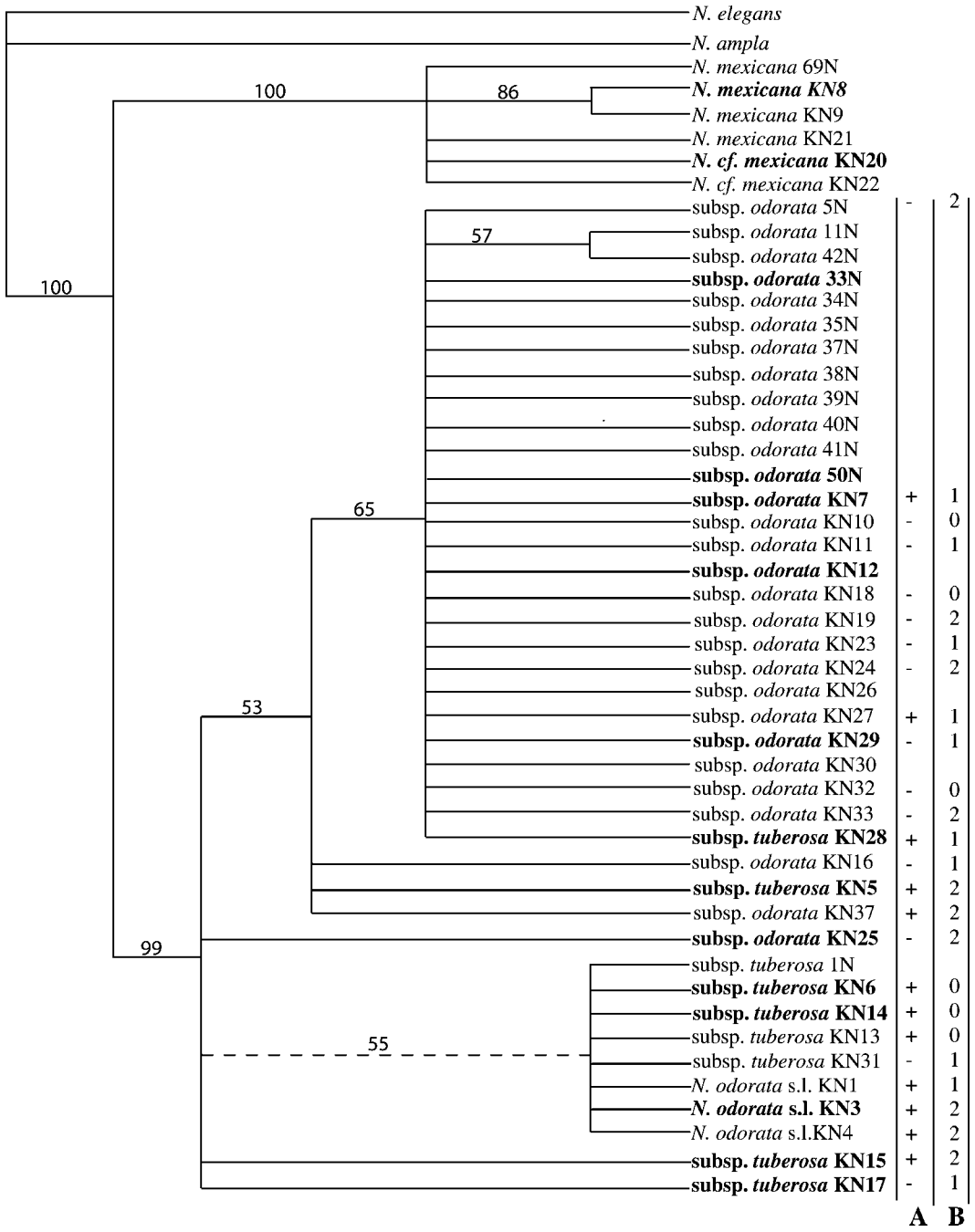

FIG. 2. Bootstrap tree based on sequences from the ITS region (length = 190, CI = 0.947, RI = 0.977). Numbers on branches indicate bootstrap values based on 500 replicates. *Nymphaea ampla* and *N. elegans* were used as outgroups. Names in bold signify samples sequenced for the *trnL-trnF* region. The clade marked by two dashed lines collapsed in the strict consensus tree. Column A indicates the presence (+) or absence (−) of petiole stripes, and column B depicts leaf blade undersurface color (0 = green, 1 = mix of green and purple, 2 = reddish-purple).

a G in subsp. *tuberosa*. These results were confirmed in this study based on a broader sampling of individuals (Table 2). The exception is sample KN7, which was identified morphologically by Borsch, Wiersema and Hellquist (Table 1) as subsp. *odorata*, but possesses the G state, implying a possible subsp. *tuberosa* chloroplast genome. The *N. cf. mexicana* sample (KN20) has the distinctive *N. mexicana* sequence.

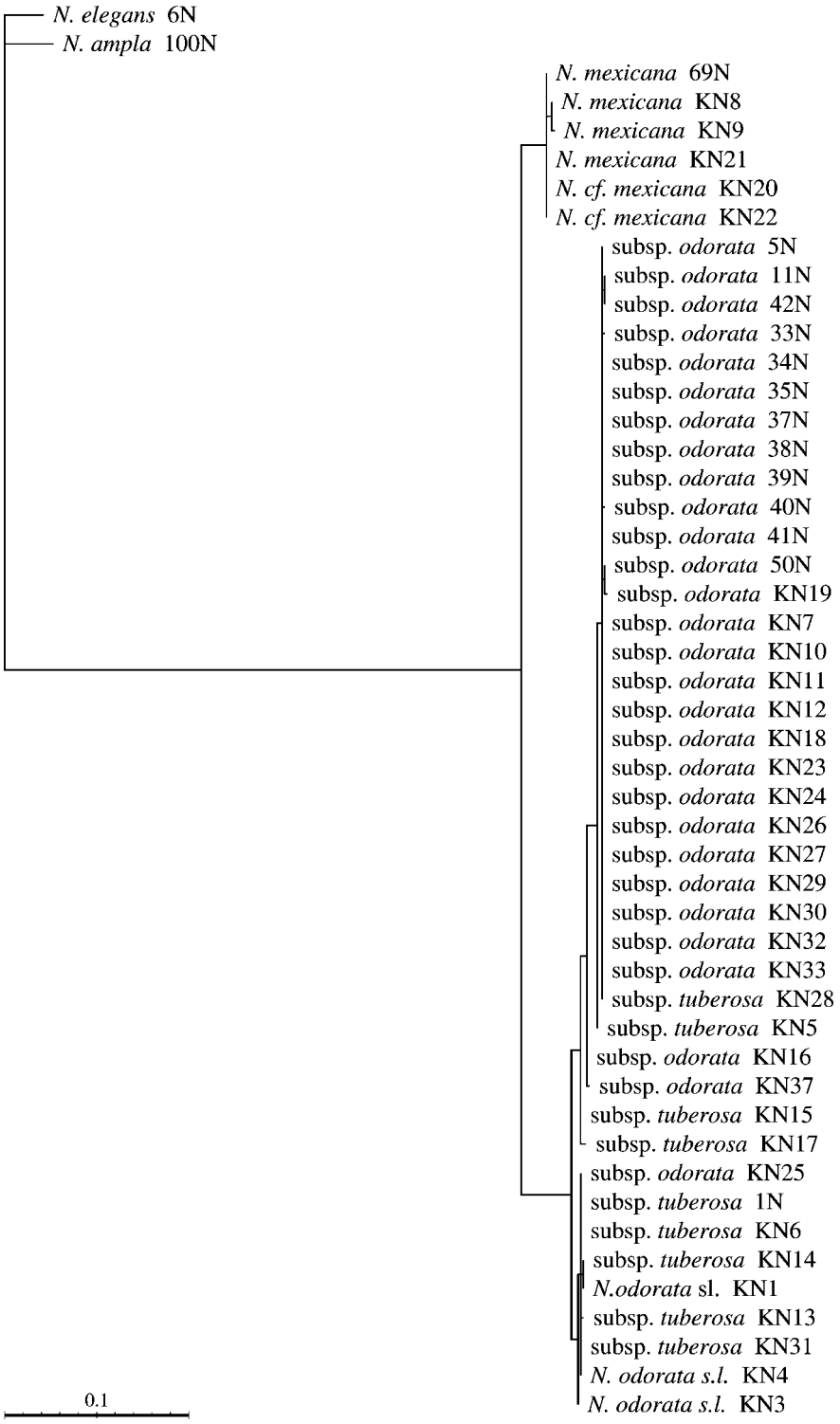

FIG. 3. Maximum likelihood tree (Lscore-1966.38181) based on ITS sequences assuming a general time reversible model (GTR) and a rate variation among sites following a gamma distribution (four categories represented by mean). *Nymphaea ampla* and *N. elegans* were used as outgroups.

TABLE 2. The identity of the one basepair molecular marker at position 396 in the *trnL-trnF* region, and genebank accession number. Samples selected were either putative hybrids between the species, representatives of each subspecies, or had unusual placement in the ITS bootstrap tree. Samples 50N, 33N, KN12, KN29, 1N, KN6 and KN14 are representatives of the subspecies and were included to check the accuracy of this nucleotide position.

KN8 from Texas, identified as *N. mexicana*, base 396 is T (AY805242). **KN20** from Florida, identified as *N. cf. mexicana*, base 396 is T (AY805241). **50N** from Georgia, identified as *N. odorata* subsp. *odorata*, base 396 is T (AY805243). **33N** from Florida, identified as *N. odorata* subsp. *odorata*, base 396 is T (AY805233). **KN12** from Louisiana, identified as *N. odorata* subsp. *odorata*, base 36 is T (AY805244). **KN29** from Vermont, identified as *N. odorata* subsp. *odorata*, base 396 is T (AY805245). *KN7* from Michigan, identified as *N. odorata* subsp. *odorata*. base 396 is G (AY805236). **KN25** from Tennessee, identified as *N. odorata* subsp. *odorata*, base 396 is T (AY805237). **1N** from New York, identified as *N. odorata* subsp. *tuberosa*, base 396 is G (AY805246). **KN3** from Michigan, identified as *N. odorata* s.l., base 396 is G (AY805248). **KN5** from Wisconsin, identified as *N. odorata* subsp. *tuberosa*, base 396 is G (AY805234). **KN6** from Manitoba, identified as *N. odorata* subsp. *tuberosa*, base 396 is G (AY805235). **KN14** from Michigan, identified as *N. odorata* subsp. *tuborosa*, base 396 is G (AY805247). **KN15** from Ohio, identified as *N. odorata* subsp. tuberosa, base 396 is G (AY805238). **KN17** from Ohio, identified as *N. odorata* subsp. *tuberosa*, base 396 is G (AY805239). **KN28** from Vermont, identified as *N. odorata* subsp. *tuberosa*, base 396 is G (AY805240).

Polymorphic Sites. Thirty polymorphic sites were found in the ITS region of *N. odorata*, of which 17 are in ITS 1, four in the 5.8S gene, and nine in ITS 2. Of the 47 samples scored, 29 were polymorphic, containing one to nine polymorphic sites. Four samples (KN7, KN17, KN31, and KN37) have the same character states for the polymorphic sites at five positions, two of these samples belong to subsp. *odorata* and two to subsp. *tuberosa* (Table 1).

SAHN clustering and PCOA analyses of the polymorphic-character matrix produce groups identical to the lineages resolved in the bootstrap tree of the ITS region. SAHN resulted in four clusters (Fig. 4) representing *N. mexicana* samples (MEX), primarily subsp. *tuberosa* (TUB), subsp. *odorata* (ODO), and subsp. *odorata* plus subsp. *tuberosa* (O/T). The MEX and TUB groups were linked at a coefficient of 0.5 and ODO and O/T were clustered at 0.6 coefficient. Similar to the bootstrap tree, the ODO group contains one sample of subsp. *tuberosa* (KN28). Sample KN5 appeared in a distinct position, associated at very low coefficient with the ODO cluster; this distinct position was apparent in other analyses (Figs. 2, 3). A PCOA analysis with a MST superimposed (Fig. 5) provides results similar to those obtained from the SAHN clustering where three groups are resolved, representing *N. mexicana*, subsp. *tuberosa*, and subsp. *odorata* (Fig. 4). Plotting the MST onto the PCOA reiterates the association between *N. mexicana* and subsp. *tuberosa*, as evident from the connections between the two groups. In addition, the MST revealed for the most part a step-wise connection from subsp. *tuberosa* samples to the subsp. *odorata* samples. The samples forming the gradation (KN15, KN25, KN37, KN17, and KN16) are the same samples found in the O/T group in the SAHN analysis; these samples appear unresolved in the bootstrap tree.

DISCUSSION

***Variability in ITS and* trnT-trnF.** There was a marked difference in sequence variability between the plastid and nuclear genomic regions studied in *N. odorata*. The chloroplast region contained only one variable site (Table 2), whereas the ITS region contained 40 variable sites between the subspecies. Borsch (2000) has found the ITS region to be extremely variable in *Nymphaea*, and was not able to completely align the ITS region throughout the genus. Numerous studies have found chloroplast spacers less variable than ITS sequences (Franzke and Hurka 2000; Stanford et al. 2000; Sang et al. 1997). Other studies, sampling different taxa, have shown the chloroplast DNA to be more informative compared to nuclear DNA (Koch et al. 1998; Martinsen et al. 2001), usually due to higher variability. However, comprehensive evaluation of gene flow benefits from a combination of evidence from both nuclear markers (which may resolve recombination) and plastid markers (which reflects uniparental inheritance) (reviewed by Linder and Rieseberg 2004).

In ITS, concerted evolution among numerous paralogues is usually assumed and many authors perform direct sequencing of PCR products for phylogeny reconstruction, treating ITS as a single locus. In this study, polymorphic sites occur, most likely caused by different alleles in heterozygous individuals resulting from hybridization. The ITS region in such individuals may be subject to recombination and concerted evolution, and ITS copies appear as different paralogues in a single array of ribosomal DNA. However, the exclusion of polymorphic sites from phylogenetic tree reconstruction in this study should minimize the effect of random copy selection. Furthermore, the exclusion of the individuals with polymorphic sites in one analysis did not influence tree topology.

***Monophyly of* N. mexicana *and* N. odorata.** Populations of the two species are resolved into two well-supported clades (99 and 100% BS) in the MP and ML trees (Figs. 2, 3) despite the potential gene flow between them. The higher variability in the ITS region compared to the *trnL-trnF* region underscores the effectiveness of ITS for *Nymphaea* systematics at both the inter- and infraspecific levels.

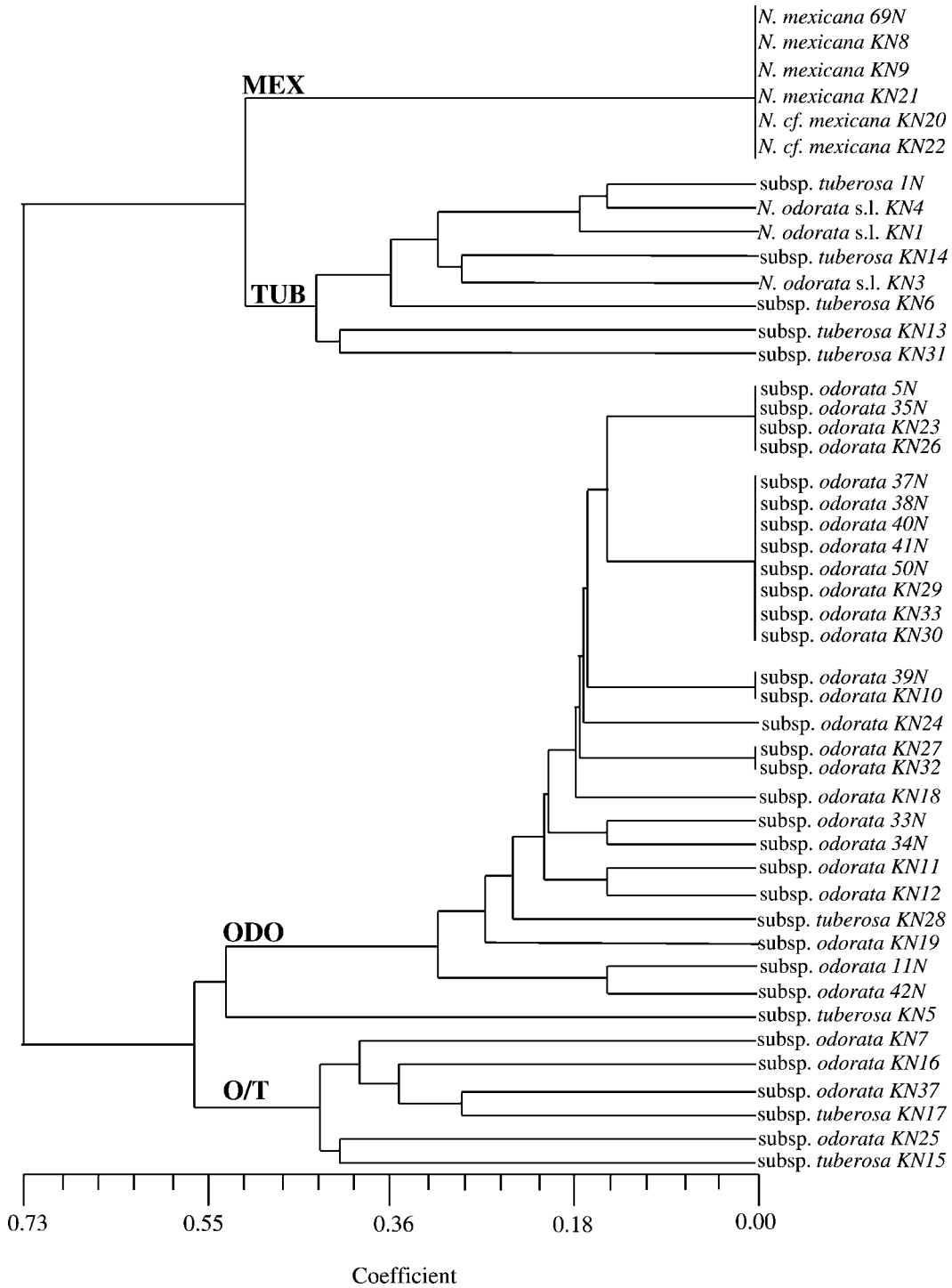

FIG. 4. Tree produced from SAHN clustering of polymorphic ITS sites. Four distinct groups are evident: MEX (*N. mexicana* and *N. cf. mexicana*), TUB (subsp. *tuberosa*), ODO (subsp. *odorata*) and O/T (six samples belonging to both subspecies).

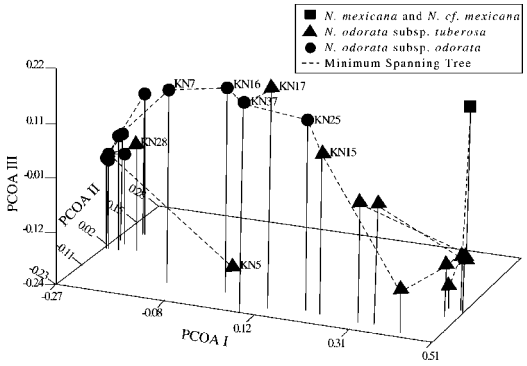

FIG. 5. Principal coordinate analysis with minimum spanning tree superimposed, generated from polymorphic ITS sites. All *N. mexicana* samples group together, and cluster closest to subsp. *tuberosa*. Several samples form a gradient from subsp. *tuberosa* to subsp. *odorata*. One sample of subsp. *tuberosa* (KN28) clusters with the main group of subsp. *odorata* samples. The proportion of total variance comprising each axis was 63.0% for axis I, 11.8% for axis II, and 5.7% for axis III.

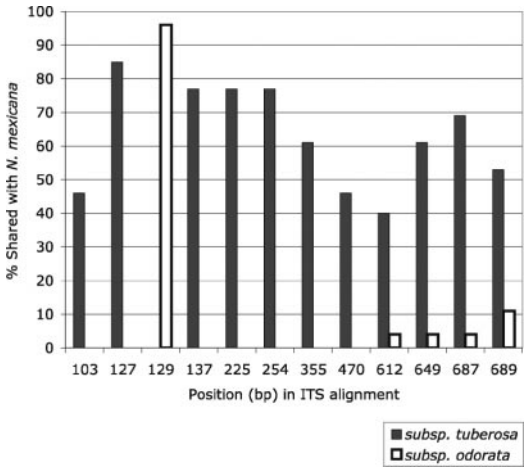

FIG. 6. Shared character states among subsp. *odorata*, subsp. *tuberosa* and *N. mexicana* in the ITS region. The percentage shared was calculated as the number of samples sharing a basepair with *N. mexicana* per the total samples of each subspecies. Most (11/12) basepairs are shared between *N. mexicana* and subsp. *tuberosa*.

***Taxonomy of* N. odorata.** Whereas the monophyly of *N. mexicana* and *N. odorata* is clear, support for segregating subsp. *odorata* and subsp. *tuberosa* is less evident. Node support in the MP analysis is low (65% and 55% BS, respectively), and branches largely reflecting these two subspecies are short in the ML tree (Figs. 2, 3). The low resolution and the appearance of individuals of one subspecies in the clade of another might be due to homoplasy, or caused by gene flow among populations of the two subspecies as suggested by the analysis of polymorphic sites (Fig. 4). These results agree with a morphological study (Woods et al. 2005) based on a principal component analysis of 26 vegetative characters, which demonstrated high variability within the species, but only partial segregation of subsp. *odorata* and subsp. *tuberosa*. Nevertheless, the one bp marker from the *trnL-trnF* region discriminates between the two subspecies (Table 2) with only a single exception, KN7, which was identified as subsp. *odorata* but possessed the G state, a synapomorphy for subsp. *tuberosa*. The placement of sample KN7 in the subsp. *odorata* clade in the ITS tree, as well as its possession of leaves with round-lobed apex and reddish purple undersurface support its identity as subsp. *odorata*. The *trnL-trnF* marker in KN7 may represent maternal gene flow from subsp. *tuberosa* or possibly a reverse mutation in this sample, reiterating problems underlying subspecies classification in *N. odorata*.

Additional support for classification at the subspecies rather than species level was provided by the analyses of ITS polymorphic sites. In all SAHN, PCOA, and MST analyses (Figs. 4, 5), clusters representing primarily subsp. *odorata* and subsp. *tuberosa* were evident. The presence in the SAHN clustering of two distinct groups (corresponding to the two subspecies) plus a group containing members of both subspecies (the O/T group) points to the possibility of gene flow between the two taxonomic entities that is partially blurring their boundaries. Gene flow may have contributed to the difficulties encountered in resolving clear-cut infraspecific lineages using morphological characters and ISSR loci (Woods et al. 2005).

Morphological characters of petiole striping and the color of the leaf blade undersurface mostly support the subspecies classification based on molecular information. These characters are mapped onto the ITS tree (Fig. 2). Mapping of these characters, cited by Wiersema and Hellquist (1997) in the *Flora of North America* (*FNA*), showed several samples (KN3, KN4, KN5, KN15, and KN37) that combine the petiole striping of subsp. *tuberosa* and the reddish-purple leaf blade undersurface characteristic of subsp. *odorata*. These samples appeared in the ITS bootstrap tree either unresolved in the subsp. *tuberosa* clade or formed a polytomy (KN5, KN37, KN16) at the base of the subsp. *odorata* clade (Figs. 2, 4). Additionally, 12 samples displayed a purplish-green leaf blade, an intermediate color between the reddish-purple of subsp. *odorata* and the green of subsp. *tuberosa*. These samples appeared throughout the *N. odorata* clade in the ITS tree (Fig. 2).

In the ITS region, there were 12 potentially parsimony-informative positions among individuals of *N. odorata*. *Nymphaea mexicana* and subsp. *tuberosa* share the same character state at 11 of these 12 positions, whereas subsp. *odorata* populations possess a different character state (Fig. 6). In addition, the MST in the PCOA analysis of polymorphic characters demonstrated linkage between *N. mexicana* and subsp. *tuberosa*

samples (Fig. 5), providing evidence for the affinity between the genomes of these two taxa. Since MST and PCOA are phenetic approaches, there is no possibility to distinguish whether these shared states are homoplasious (i.e., being the result of parallel or convergent substitutions), or symplesiomorphic and thus retained from a common ancestor of *N. mexicana* and *N. odorata* subsp. *tuberosa*. In contrast, in the *trnL* intron, *N. mexicana* and subsp. *odorata* share the character state ''T,'' whereas subsp. *tuberosa* has character state ''G.'' This position was found to be among the most homoplastic positions in basal angiosperms (Borsch et al. 2003), where it has repeatedly changed from G to T, (with G the plesiomorphic state). However, within *Nymphaea*, the state is usually a G, mutating into a T only in *N. mexicana* and *N. odorata* subsp. *odorata*.

The evolution of this variable position in the *trnL* intron should be examined in a phylogenetic context. Various hypotheses exist (Fig. 7) based on the position of *N. odorata* and *N. mexicana* within the temperate clade (Borsch 2000). The first-branching position of *N. mexicana* alone (Fig. 7a), however, appears more feasible when morphology is considered and thus these alternatives will be evaluated first. In this case, the most parsimonious hypothesis, involving two mutational steps, would require parallel G→T only in *N. mexicana* and subsp. *odorata*, thereby the T is a derived state in both taxa. Additionally, there are two explanations of three steps each based on a G→T mutation in the common ancestor of the temperate clade: (1) a T→G reversal after the split of *N. mexicana* and a repeated G→T substitution in subsp. *odorata*, and (2) two independent T→G reversals in subsp. *tuberosa* and the core of the temperate clade. Considering this longer pathway, the shared T state of subsp. *odorata* with *N. mexicana* would be explained by either reversal or plesiomorphy. Other alternatives, but less likely considering morphology, are based on *N. odorata* and *N. mexicana* being sisters (Fig. 7b,) and include two most parsimonious reconstructions of two steps: (1) G→T in the common ancestor of the *N. odorata*-*N. mexicana* clade and a T→G reversal in subsp. *tuberosa*, where locally subsp. *tuberosa* would have the derived state or, as in Fig. 2A, (2) independent and parallel G→T in *N. mexicana* and subsp. *odorata*, giving them the derived state. Although additional information is needed for a conclusive assessment of the cp haplotype evolution, the most parsimonious scenarios point to either parallel evolution or possible ancient introgression. Nevertheless, the result of two geographically more or less separated cp haplotypes is an important finding of this study.

Compared to the distribution of cp haplotypes, the distribution of ITS genotypes reflects a more complex picture. Here a number of substitutions are obviously shared between *N. mexicana* and subsp. *tuberosa*. This

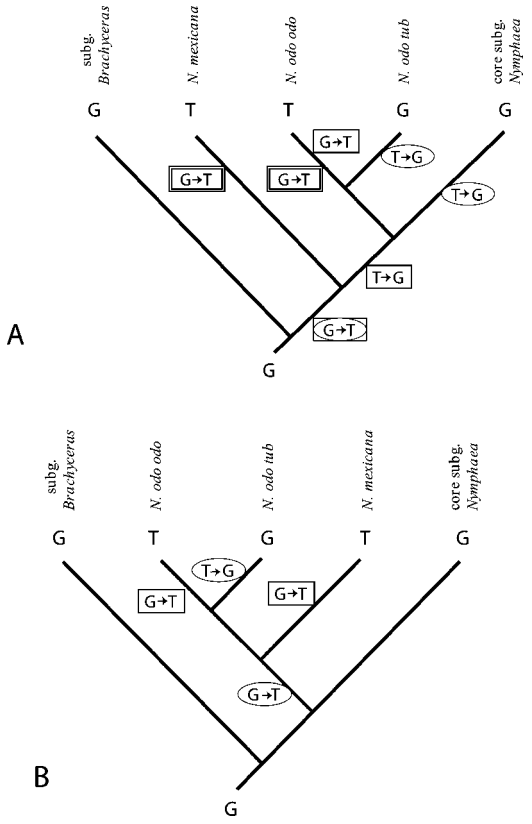

FIG. 7. Distribution of the two states (G and T) for the variable *trnT-trnF* site on two alternative hypotheses for the phylogenetic relationship in section *Nymphaea* following Borsch (2000). The ''G'' state is plesiomorphic in *Nymphaea*. Constraint tree A (probable): The most parsimonious scenario is that *N. mexicana* and *N. odorata* subsp. *odorata* aquired the state ''T'' in parallel (double boxes). Thus, subsp. *odorata* would have a derived cp haplotype. Two other alternatives involving one additional step, and thus less parsimonius, are also indicated (single boxes or ovals). Constraint tree B (less probable): There are two equally parsimonious scenarios. If a ''T'' was gained in parallel (boxes), subsp. *odorata* would have a derived cp haplotype. If the ''G'' in subsp. *tuberosa* results from a reversal (ovals), subsp. *tuberosa* would have the derived cp haplotype. Scenarios with three mutational steps not shown. ''G'' is plesiomorphic in *Nymphaea*.

should not be seen as evidence of a common ancestry exclusive of subsp. *odorata*, which would be contradicted by morphology. These shared character states could either involve a considerable amount of homoplasy or ancient gene flow from an ancestor of *N. mexicana* to an ancestor of subsp. *tuberosa*, followed by concerted evolution in the direction of one parent (Wendel et al. 1995). This introgression would have taken place with *N. mexicana* ancestor as the paternal species since the plastid genome types of *N. odorata* have obviously been maintained. Ancient hybridization is proposed based on the lack of polymorphisms in the shared sites between *N. mexicana* and subsp. *tuberosa*, indicating that

concerted evolution was able to reach homogeneity. This is in contrast to the proposed hybridization between the two subspecies of *N. odorata*, where the presence of polymorphic sites indicates relatively recent hybridization events.

A third possibility is that these shared characters are plesiomorphic in subg. *Nymphaea*, cannot be eliminated without additional study. From this perspective, information from the life history of subsp. *tuberosa* may provide further explanation for the higher proportions of synapomorphies among individuals of subsp. *odorata* in ITS. In populations of subsp. *odorata*, seed set is almost always found, whereas in populations of subsp. *tuberosa* seed set is rare and asexual propagation seems to predominate (J. Wiersema unpubl. data). Therefore, subspecies *tuberosa* is behaving more like a clone, whereas subsp. *odorata* generates and maintains high levels of genetic diversity via outcrossing. It is possible that increased time between sexual generations in subsp. *tuberosa* may decrease the chance of mutations to become fixed, making it difficult for this subspecies to accumulate unique mutations compared to a predominantly sexually reproducing organism such as in subsp. *odorata*.

***Evidence of Hybridization in* N. odorata.** Species that are perennial, outcrossing breeders and have the ability to asexually reproduce are the most likely to hybridize (Rieseberg 1997). *Nymphaea odorata* possesses all of these characteristics. Hybrids between *N. mexicana* and *N. odorata* subsp. *odorata* and between the two subspecies of *N. odorata* have been reported (Wiersema and Hellquist 1997). The interspecific hybrids are sterile (Wiersema and Hellquist 1997). However, the suspected intraspecific hybrids in *N. odorata* are fertile, and can be difficult to identify based on morphology alone (Wiersema and Hellquist 1997). Borsch and Wiersema (pers. comm.) originally identified three samples (KN1, KN3, and KN4) as ''putative hybrids'' because they combined morphological characteristics of both subspecies (Fig. 2) that were sympatric in a Michigan population. However, analyses of ITS sequences show these samples nested within the subsp. *tuberosa* clade (Figs. 2, 3), and were part of the subsp. *tuberosa* cluster in the phenetic analyses of the polymorphic sites (Figs. 4, 5). Therefore, either the ITS region was concerted towards the subsp. *tuberosa* type, or their morphological intermediacy reflects either plasticity or an extreme case of overlap in morphologies between the two subspecies. Hybrids can have a unique effect in phylogenetic analyses. They can form polytomies between sister and non-sister taxa, or result in a basal trichotomy showing no relationship among ancestral taxa (Humphries 1983). In the ITS bootstrap tree, five samples (KN15, KN16, KN17, KN25, KN37) did not group with either subspecies (Fig. 2). In the strict consensus tree (indicated in Fig. 2), these samples, along with others from subsp. *tuberosa*, formed a polytomy at the base of the subsp. *odorata* clade. These five samples appeared in a grade leading to the subsp. *odorata* lineage in the ML (Fig. 3). The combination of morphological characters from both subsp. *odorata* and subsp. *tuberosa* in these samples and their position in the MP and ML trees may imply that they are intersubspecific hybrids. These five samples possess six to nine polymorphisms in the ITS region, further supporting their potential hybrid origin. To address the potential effect of these putative hybrids on the tree topology, we excluded them and reanalyzed the data. The primarily subsp. *odorata* clade was recovered with increased bootstrap support (65% to 98%). However, the subsp. *tuberosa* samples formed a polytomy with the subsp. *odorata* clade (trees not shown). Therefore, the overall topology of the MP tree was not altered upon removal of the hybrids. These results are similar to McDade's (1990, 1992) finding in which no change in tree topology was detected after removal of hybrid taxa. McDade (1990) indicated that hybrids in phylogenetic analyses do not lead to unresolved cladographs with rampant homoplasy, instead they are placed as a basal lineage to the clade that includes its most derived parent (McDade 1990). At the same time, neither phylogenetic nor phenetic analyses can provide information on the specific evolutionary history of the hybrids (McDade 1997). Although our phylogenetic analysis concurs with this conclusion, our PCOA/MST phenetic analysis may provide insight into the pattern of hybridization in *N. odorata*.

In this study, samples KN7, KN17, KN31, and KN37 share exactly the same character states in five polymorphic ITS sites. Although KN17 and KN37 are part of the ML grade mentioned above, KN7 was nested in the subsp. *odorata* clade and KN31 in the subsp. *tuberosa* clade. However, in the analysis of the polymorphic sites, KN7 and KN37 were not closely associated with either subspecies in the PCOA and were part of the series of individuals that link the two subspecies in the MST (Fig. 5). Polymorphisms point to the hybrid origin of samples KN7, KN17, KN31, and KN37.

Considering the occurrence and pattern of polymorphic sites, it appears that concerted evolution has reached different degrees of homogenization in the directions of the two subspecies. Complete concerted evolution in the ITS region has been recently questioned by the presence of polymorphisms (Sang et al. 1995, Wendel et al. 1995, Franzke and Mummenhoff 1999, Barkman and Simpson 2002). This is evident in *N. odorata*.

Polymorphic sites in the ITS region can also provide insight into relationships between hybrids and parental taxa. Franzke and Mummenhoff (1999) suggested that recent hybrids contain polymorphic sites in the ITS region that were shared with the parents. In this

study, polymorphisms were analyzed by both SAHN clustering, PCOA and MST analyses. The SAHN clustering analysis is based on average distance between samples, whereas the MST analysis depicts relationships between nearest neighbors. Consequently, MST is a better indicator of direct relationships between individual samples of *N. odorata*. The segregation in SAHN of three clusters (Fig. 4) representing individuals identified as subsp. *odorata*, subsp. *tuberosa* and those recognized as potential hybrids points to the effective use of polymorphic sites in this study. The potential pattern of gene flow was further elucidated in the PCOA and MST analyses (Fig. 5). MST reveals more connections within each subspecies than between them, implying more frequent gene flow within than between the two subspecies. In addition, MST also points to potential gene flow among populations of the two subspecies where subsp. *odorata* and subsp. *tuberosa* haplotypes were connected through a gradation of samples KN7, KN15, KN16, KN17, KN25, and KN37. Those haplotypes did not group with either subspecies in the MP and ML analysis based on all potentially parsimony variable sites or, in the case of KN7, was nested within the subsp. *odorata* clade in the MP and ML (Figs. 2, 3), possibly due to an advanced stage in ITS concerted evolution. The emergence of KN5 in an isolated position in the SAHN clustering and PCOA/MST analyses of polymorphic sites (Figs. 4, 5) may indicate excessive divergence in that Wisconsin population; more samples from that area would need to be sampled, particularly in light of morphological intermediacy between the two subspecies.

The geographic distribution and morphological characteristics of suspected hybrids in *N. odorata* may provide further insights into patterns of hybridization. Samples KN15 and KN17 from Ohio are identified as subsp. *tuberosa*, but display mixed morphologies from the two subspecies and possess a unique flower color. Samples KN16, KN25, and KN37, found in Delaware, Tennessee, and Virginia, respectively, and identified as subsp. *odorata*, also possess some unique morphological characteristics. Conard (1905) recognized the increased leaf and flower size of KN16 and classified this morphotype as var. *gigantea* Tricker. The area where these samples are distributed may represent a possible hybrid zone between subsp. *odorata* and subsp. *tuberosa* in central eastern United States. Sympatry in this region might have been enhanced by the advancement of glaciers in North America, which may have contributed to increased hybridization due to secondary contact (Stebbins 1985). In other regions where gene flow between the two subspecies is less likely to occur, divergence continued at the morphological level and near concerted evolution was achieved in the ITS region. However, additional studies at the population level using more markers are still needed to evaluate the pattern of hybridization and speciation in *N. odorata*.

Information from ITS and *trnL-trnF* as well as morphology and ISSR markers (Woods et al. 2005) strongly support the recognition of a single species, *N. odorata*, to encompass two subspecies, subsp. *odorata* and subsp. *tuberosa*. However, the segregation of the populations into two well-defined groups, although detectable, it is not clear-cut. Such degrees of differentiation are not unexpected at the infraspecific level (Speer et al. 1999a, b). Furthermore, the situation in *N. odorata* could be attributed to the relatively recent differentiation of the two taxonomic entities and/or to hybridization events among some of their populations.

ACKNOWLEDGEMENTS. This research was supported by grants from the Virginia Academy of Sciences, International Water Lily Society in 2000 and 2001, Virginia Tech Biology Department, and the Virginia Tech Graduate Research Development Grant. We thank Susana Neves for her lab and collecting assistance, Dietmar Quandt for assistance with the ML analysis, and Jim Woods for assistance with graphics. Special thanks to C. Barre Hellquist for his help with the New England and Upper Great Lakes collections, Alejandro Novelo for the collections from Mexico, Don Les for providing population samples, and Charles Horn for some South Carolina collections. We also thank Jeff Rooke, Fred Rich, Bill Summers, Volker Wilde, Bill Williams, and Thomas Wieboldt for field assistance.

LITERATURE CITED

AKAIKE, H. 1974. A new look at the statistical identification model. *IEEE Transactions on Automatic Control* 19: 716–723.

BARKMAN, T. J. and B. B. SIMPSON. 2002. Hybrid origin and parentage of *Dendrochilum acuiferum* (Orchidaceae) inferred in a phylogenetic context using nuclear and plastid DNA sequence data. *Systematic Botany* 27: 209–220.

BORSCH, T. 2000. *Phylogeny and evolution of the genus* Nymphaea. PhD thesis, University of Bonn.

———, K. W. HILU, V. WILDE, and W. BARTHLOTT. 1998. Inferring the phylogeny of *Nymphaea*: evidence from the *trnT-trnF* region of the chloroplast genome. *American Journal of Botany* 85(6, supplement): 116–117.

BORSCH, T., K. W. HILU, D. QUANDT, V. WILDE, C. NEINHUIS, and W. BARTHLOTT. 2003. Non-coding, fast evolving plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *Journal of Evolutionary Biology* 16: 558–567.

CONARD, H. S. 1905. *The waterlilies: a monograph of the genus* Nymphaea. Washington DC: The Carnegie Institution of Washington.

CORRELL, D. S. and H. B. CORRELL. 1975. *Aquatic and wetland plants of southwestern United States.* Stanford: Stanford University Press.

CROW, G. E. and C. B. HELLQUIST. 2000. *Aquatic and wetland plants of northeastern North America.* Madison: The University of Wisconsin Press.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.

FRANZKE, A. and H. HURKA. 2000. Molecular systematics and biogeography of the *Cardamine pratensis* complex (Brassicaceae). *Plant Systematics and Evolution* 224: 213–234.

——— and K. MUMMENHOFF. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae)—conversion of nuclear ribosomal ITS sequences in statu nascendi. *Theoretical and Applied Genetics* 98: 831–834.

HOLMGREN, P. K., N. H. HOLMGREN, and L. C. BARNETT. 1990. *Index herbariorum*, edition 8, Part I. *The herbaria of the world*. Bronx: New York Botanical Garden Press.

GLEASON, H. A. and A. CRONQUIST. 1991. *Manual of vascular plants of northeastern United States and adjacent Canada*, ed. 2. Bronx: The New York Botanical Garden Press.

HUMPHRIES, C. J. 1983. Primary data in hybrid analysis. Pp. 89–103 in *Advances in cladistics: Proceedings of the Second Meeting of the Willi Hennig Society*, eds. N. I. Platnick and V. A. Funk. New York: Columbia University Press.

KOCH, M., K. MUMMENHOFF, and H. HURKA. 1998. Molecular biogeography and evolution of the *Microthlaspi perfoliatum* s.l. polyploid complex (Brassicaceae): chloroplast DNA and nuclear ribosomal DNA restriction site variation. *Canadian Journal of Botany* 76: 382–396.

LIANG, H. P. and K. W. HILU. 1996. Application of the *matK* gene sequences to grass systematics. *Canadian Journal of Botany* 74: 125–134.

LINDER, C. R. and L. H. RIESEBERG. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91: 1700–1708.

MARTINSEN, G. D., T. G. WHITHAM, R. J. TUREK, and P. KEIM. 2001. Hybrid populations selectively filter gene introgression between species. *Evolution* 55: 325–1335.

MCDADE, L. A. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44: 1685–1700.

———. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.

———. 1995. Hybridization and Phylogenetics. Pp. 305–331 in *Experimental and molecular approaches to plant biosystematics*, eds. P. C. Hoch and A. G. Stephenson. St. Louis: Missouri Botanical Garden Press.

———. 1997. Hybrids and phylogenetic systematics III. Comparison with distance methods. *Systematic Botany* 22: 669–683.

MCGREGOR, R. L., T. M. BARKLEY, R. E. BROOKS, and E. K. SCHOFIELD. 1986. *Flora of the Great Plains.* Lawrence: University Press of Kansas.

MÜLLER, J. 2002. QuickAlign 1.5.7 beta version. University of Bonn, Germany. http://www.uni-bonn.de/downloads/QA.

PATTI, F. P. 2002. WinModeltest front-end v4b. Program distributed by the author, Harvard.

POSADA, D. and K. A. CRANDALL. 2001. Modeltest v3.06: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

RIESEBERG, L. H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28: 359–389.

ROHLF, F. J. 1998. NTSYS-PC numerical taxonomy and multivariate analysis systems, Version 2.02k. Applied Biostatistics Inc., Setauket, NY.

SANG, T., D. J. CRAWFORD, and T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proceedings of the National Academy of Science* 92: 6813–6817.

———, ———, and ———. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120–1136.

SOKAL, R. R. and C. D. MICHENER. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38: 1409–1438.

SPEER, B., K. W. HILU, and C. WERTH. 1999a. Relationships between two infraspecific taxa of *Pteridium aquilinum* (Dennstaedtiaceae). I. Morphological evidence. *Systematic Botany* 23: 305–312.

———, ———, and ———. 1999b. Relationships Between Two Infraspecific Taxa of *Pteridium aquilinum* (Dennstaedtiaceae). II. Isozyme evidence. *Systematic Botany* 23: 305–312.

STANFORD, A. M., R. HARDEN, and C. R. PARKS. 2000. Phylogeny and Biogeography of *Juglans* (Juglandaceae) based on *matK* and ITS sequence data. *American Journal of Botany* 87: 872–882.

STEBBINS, G. L. 1985. Polyploidy, hybridization and the invasion of new habitats. *Annals of the Missouri Botanical Garden* 72: 824–832.

SWOFFORD, D. L. 2000. PAUP*-Phylogenetic analysis using parsimony (* and other methods). Version 4.0b10. Sunderland: Sinauer Associates.

TABERLET, P., L. GIELLY, G. PAUTOU, and J. BOUVET. 1991. Universal primers for the amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

VOSS, E. G. 1985. *Michigan flora Part II (Saururaceae Cornaceae).* Ann Arbor: Cranbrook Institute of Science and University of Michigan Herbarium.

WARD, D. B. 1977. Keys to the flora of Florida—4, *Nymphaea* (Nymphaeaceae). *Phytologia* 37: 443–448.

WENDEL, J. F., A. SCHNABEL, and T. SEELANAN. 1995. An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. *Molecular Phylogenetics and Evolution* 4: 298–313.

WHITE, T. J., T. BRUNS, S. LEE, and J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in *PCR protocols: a guide to methods and applications*, eds. D. G. M. Innis, J. Sninsky, and T. White. San Diego: Academic Press.

WIERSEMA, J. H. and C. B. HELLQUIST. 1994. Nomenclatural notes in Nymphaeaceae for the North American Flora. *Rhodora* 96: 170–178.

——— and ———. 1997. *Nymphaea* Linnaeus. Pp. 71–77 in *Flora of North America North of Mexico* vol. 3., eds. Flora of North America Editorial Committee. New York: Oxford University Press.

WOODS, K., K. W. HILU, J. H. WIERSEMA, and T. BORSCH. 2005. Pattern of variation and systematics of *Nymphaea odorata*: I. Evidence from morphology and inter-simple sequence repeats (ISSRs). *Systematic Botany* 30: 471–480.

APPENDIX 1. Ambiguous basepairs present in the ITS analyses. Samples were scored for each polymorphic position as present (1) or absent (0) for each basepair.

69N	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
KN8	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
KN9	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
KN21	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
KN20	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
KN22	1 0 0 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 1 0
005N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
011N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 1 1 0 1 0
33N	1 0 0 0 1 0 1 0 1 0 1 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
034N	1 0 0 0 1 0 1 0 1 1 1 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
035N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
037N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
038N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
039N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 0 1 0
040N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
041N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
042N	1 0 0 0 1 0 1 0 1 1 1 1 0 1 0 1 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 1 1 0 1 0
50N	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
KN7	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 1 1 0 0 0 1 1 1 1 1 1 1 1 1 1 0 0 1 1 1
KN10	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 0 1 0
KN11	1 0 0 1 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
KN12	1 1 0 1 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
KN16	1 0 0 0 1 0 1 0 1 1 0 1 1 1 0 1 0 0 0 1 0 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 0 0 0 1 1 1
KN18	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 0 1 0 0 1 0 0 1 1 1 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0
KN19	1 0 0 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 1 0 0 1 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0